NASA Technical Memorandum 86413

FIT: Computer Program That Interactively Determines Polynomial Equations for Data Which Are a Function of Two Independent Variables

P. Douglas Arbuckle, Steven M. Sliwa,
Marie-Louise Roy, and Sherwood H. Tiffany

OCTOBER 1985

NASA

FL 2827
WSMC/PMET TECHNICAL LIBRARY
VANDENBERG AFB CA 93437-6021

NASA Technical Memorandum 86413

FIT: Computer Program That Interactively Determines Polynomial Equations for Data Which Are a Function of Two Independent Variables

P. Douglas Arbuckle and Steven M. Sliwa
Langley Research Center
Hampton, Virginia

Marie-Louise Roy
PRC Kentron, Inc.
Hampton, Virginia

Sherwood H. Tiffany
Langley Research Center
Hampton, Virginia

National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1985

Summary

A computer program for interactively developing least-squares polynomial equations to fit user-supplied data is described. The program is characterized by the ability to compute the polynomial equations of a "surface" fit through data that are a function of two independent variables. The program utilizes the Langley Research Center graphics packages to display polynomial equation curves and data points, facilitating a qualitative evaluation of the effectiveness of the fit. An explanation of the fundamental principles and features of the program, as well as sample input and corresponding output, are included.

Introduction

In science and engineering research, it is often valuable to calculate the equation of a "smooth" curve or surface through data points. If a useful program is to be developed to exploit the capability of computing an equation to fit data, the program must include graphic output. This is because many algorithms will yield equations which pass close to or through the data points, but the resulting curve or surface may not have realistic shape characteristics. Ideally, most engineers would prefer an interactive program with computer graphics for fitting a curve or surface to data. The resulting ability to model experimental or tabular data with one analytic equation, as opposed to looking up values in a table and interpolating these values, can be extremely valuable.

This report describes a computer program, FIT, which calculates the polynomial equation of a least-squares fit through user-supplied data (which are a function of either one or two independent variables) and which allows a user at an interactive terminal to: "control" the curve or surface fit, edit the data input within the program, view a plot of the polynomial fit through the data, and save the plotting parameters of the last plot viewed for future use. Various plotting formats may be selected to aid the user in visualizing and evaluating the quality of a surface fit. The interactive modules of the program are menu or prompt driven with "HELP" options for user-friendly operation. Informative output files are created which may be used for detailed evaluation of the quality of a curve or surface fit to the input data.

This program has been designed primarily for ease of use and speed of computation, since these factors are desirable for interactive software. This philosophy led to the use of least-squares polynomial equations for fitting the data as opposed to use of more sophisticated (and more complex to use) techniques such as splines. For the "real world" data to which the authors typically apply this program, the choice of data fitting technique has been appropriate.

The FIT computer program is available as LAR-13457 from COSMIC, 112 Barrow Hall, University of Georgia, Athens, GA 30602.

Symbols

A	polynomial coefficients for a given surface fit of an $(E + 1) \times M$ matrix
B	coefficients for a given polynomial fit of an $M \times L$ matrix
E	user-selected order of polynomial fit in Z, $E < L$
L	number of sets of dependent-variable values Y; $L = 1$ means that Y is a function of one independent variable X and $L > 1$ means that Y is a function of two independent variables X and Z
M	number of polynomial coefficients determined; $M - 1$ is the order of polynomial fit in X
N	number of X-values
X	"first" independent variable
Y	dependent variable, $Y = f(X)$ or $Y = f(X,Z)$
Z	"second" independent variable

Description of Program Features and Principles

General

A schematic of the FIT program hierarchy is shown in figure 1. FIT may be initialized either interactively or from the data on a file called DATIN (see appendix A), which is written at the end of every FIT session. After initialization, the program proceeds to the option menu. From this menu, the user may select any of the appropriate FIT options (discussed in detail later in this paper). As each option is completed, the user is returned to the option menu. When the user terminates the program, output files containing the coefficients, informative output (DATOUT), and initialization information (DATIN) are written based on the program parameters at termination.

Determining Polynomial Equations to Fit Data

FIT uses a Langley FORTRAN subroutine, LSQQRA (from Langley Central Scientific Computing Complex Document N-3a), to compute the coefficients of a least-squares polynomial of specified order $(M - 1)$ which best fits the input data. The subroutine accepts data in the form of L sets of dependent-variable values Y associated with a specified number of values N of an independent variable X. The coefficients of the polynomial are determined from the weighted least-squares solution found with the methods of references 1 and 2. Also, the residuals and their sums of squares are computed.

For data which describe a curve, the computed polynomial coefficients determine the equation

$$f(X) = B_0 + B_1X + B_2X^2 + \ldots + B_iX^i$$

where i is equal to $M - 1$ and X varies from X_1 to X_N (input data values). For surface data in the functional form $f(X, Z)$, the following technique is used to compute the coefficients that determine the polynomial equation for f(X,Z):

1. For each of L curves, the coefficients of a polynomial of order $M - 1$ are computed.

$$f(X,Z_\ell) = B_0 + B_1X + B_2X^2 + \ldots + B_iX^i \qquad (\ell = 1, 2, \ldots, L)$$

where i is equal to $M - 1$ and X varies from X_1 to X_N (input data values).

2. The coefficients B_i (for each of L curves) then become "dependent-variable values" of M curves. For each of these M curves, the coefficients of a polynomial of Eth order (less than L) are then computed.

$$B_i(Z) = A_{0i} + A_{1i}Z + A_{2i}Z^2 + \ldots + A_{ji}Z^j \qquad (i = 0, 1, \ldots, M - 1)$$

where j is equal to E and Z varies from Z_1 to Z_L (input data values).

The equation for $f(X,Z)$ can now be expressed in a polynomial equation that is a function of both X and Z and that requires knowledge of only the coefficients A_{ji}.

$$f(X,Z) = B_0(A_{j0},Z) + B_1(A_{j1},Z)\,X + B_2(A_{j2},Z)\,X^2 + \ldots + B_i(A_{ji},Z)\,X^i$$

where j varies from 0 to E for each B term and i equals $M - 1$; X and Z are user-input data values. The potential pitfall of this technique occurs if a "good" least-squares fit cannot be obtained for the coefficients B_i (step 2 above). Although this does not often occur, the FIT user must be alert to it. This is because the plot of X versus $f(X,Z)$ for a surface fit is based on polynomial equations which use the calculated coefficients B (from step 1) above and does not reflect inaccuracies that may have occurred in step 2, wherein the B coefficients are fit in a least-squares fashion to a polynomial in Z of order E. For this reason, alternative plotting options are available when computing polynomial fits to surface data to allow the user to visualize the suitability of the surface fit.

Special Program Considerations

The tolerable range (difference between the maximum and minimum values) in the user-input data (X, Y, and Z) is limited for this program because of numerical considerations. In general, attempts should be made to keep the range of the input data greater than 1 percent of its average absolute value. For example, if the average absolute value of input X-values is 100, the difference between the maximum X-value and minimum X-value should be at least 1. This test should be applied to the Y- and Z-values as well. For polynomials of about order 6 or more, the range of the input data should exceed 2 percent of its average absolute value.

The specified order of the polynomial fit should be the minimum order required to achieve a smooth and realistic fit. This also alleviates potential numerical difficulties and yields less bulky polynomial equations. The user will find that higher order polynomials give polynomial fits that pass through (or close to) all the data points but which may not be practical for engineering use. Polynomials which are of even order generally provide fits which have slopes of opposite signs at the endpoints. Polynomials of odd order generally provide fits which have endpoint slopes of the same sign. In many cases, intuition, experimentation, and persistence will be required to realize the "best" polynomial fit.

The user can attempt to constrain the polynomial fit to go through various points by responding to the appropriate prompts. The program "weights" the points when the least-squares polynomial fit is computed. Although this does not guarantee that the curve or surface fit will pass through the desired points, the polynomial fit will be closer to that point than if it had not been weighted. Satisfactory results are generally achieved when pairs of points are used to attempt to constrain various parts of the curve or surface fit. Also, for maximum effect, as few points as possible should be weighted.

The "best" polynomial fit to the input data is generally determined from a subjective judgment made by the user. The user usually knows what the fit through the data should look like. By changing the order of the fit, the user can manipulate the curve to the shape desired. Further fine-tuning of the shape is possible with judicious use of the option to attempt to constrain the polynomial fit. In some cases, the user may weight all but a few "bad" data points to achieve the desired fit. The polynomial fit does not determine the only possible curve or surface fit but merely provides the equations when a fit is achieved that is satisfactory to the user.

The surface fit capability of FIT is referred to as the "cross-plot option" both in this report and the interactive prompts within the program. This terminology is introduced to remind the user that although a surface fit polynomial equation may be obtained with this program, the technique used is to fit the coefficients of polynomial fits in X "across" Z.

Finally, the program contains the following arbitrary restrictions:

1. The number of sets L of Y-values may not exceed 11.

2. The maximum order $M - 1$ of the least-squares polynomial fit may not exceed 11.

3. The maximum number of X-values N may not exceed 100.

Input Data Processing

If required, scale factors are computed by FIT for the

input X- and Y-values such that no value has an absolute value less than 1 or an absolute value greater than 10. Although the methods of references 1 and 2 are numerically robust, this numerical conditioning helps to prevent algorithm failures because of numerically ill-conditioned problems. The scale factors also help alleviate numerical difficulties when the degree (order) of the polynomial and the values of X are large.

Initialization of Program

The first task FIT performs is to check for the existence in the local file environment of the file DATIN, which contains the data required to initialize the program. If file DATIN exists, it is read. No extensive effort is made to isolate errors since FIT assumes the DATIN file has been created by a previous FIT run. However, the following error checks are performed:

1. The setup for attempted constraint of the curve or surface fit (see section entitled *Special Program Considerations*) is checked for correctness. If any problems are found, attempted constraint is removed.
2. If the cross-plot option (surface fit) is activated, the Z-values are checked for duplicates. If any are found they are flagged.

Since it is impossible for the above conditions to exist when the DATIN file is written at the end of the program, these errors tend to occur when the DATIN file has been incorrectly edited by the user and is not in the proper format for the program.

If no DATIN file is present, an interactive subroutine is used to initialize the program. Like all other interactive subroutines in FIT, for each yes/no prompt a default response is provided. Denoted by "RET=," the default response is what the program will assume if the "return" key is depressed. A "HELP" response is provided for some of the more complicated prompts to allow the user to get additional information before responding.

For functions of two independent variables, the program asks for input in the form $X,f(X,Z)$ for constant values of Z. (See appendix B.) This causes $f(X,Z)$ to be fit as a function of X for a set of constant Z-values. (See section entitled *Determining Polynomial Equations to Fit Data*.) This is what is meant in appendix C by the prompt "ENTER NUMBER OF DEPENDENT DATA POINTS (F(X,Z) VALUES) PER X-VALUE." As an example, if there are five $X,f(X,Z)$ curves of constant Z, the user should input "5" in response to the above question.

Input data to be fit may be either entered interactively or read from any local file which has been previously prepared in the format shown in appendix B. The user is then prompted to enter a plot title, axis labels, and order of polynomial fit. If zero is entered in response to this last prompt the program plots the input data without a polynomial fit; this is a useful option for checking data entry. Next, the user is asked whether constraint of the curve or surface fit will be attempted. If the response is affirmative, prompts lead the user to select the appropriate points at which constraint will be attempted.

The next series of prompts concern how the axes are to be scaled. In the automatic mode, the user allows the program to select the axis scale. In the manual mode, the user selects the starting point and the endpoint of the axis and the number of major tic marks (where numbers will be marked on the axis) for the axis. To assist the user in selecting the proper starting point and endpoint for the axis, the maximum and minimum values of the data to be plotted are displayed.

Various options are provided for "customizing" the plot for various purposes and applications. Options are provided for the following:

1. *Data reference lines.* These are reference lines drawn at user-specified values. Data reference lines may be drawn on either axis individually or on both at the same time. This option disables the zero line and the plot grid option described below.
2. *Zero lines.* These are reference lines drawn at the axis zero. Zero lines may be drawn on either axis individually or on both at the same time. This option disables the plot grid option described below.
3. *Plot grid.* This option draws a plot grid based on the tic mark locations. This option may only be activated if there are no zero lines or data reference lines.
4. *Second X-axis.* This option draws a second horizontal axis below the first. It allows for dual-scale plots, since the user need only provide the conversion factor between the two horizontal axis scales. (For example, one axis is scaled in U.S. Customary Units while the other axis is scaled in International System of Units.) This option is disabled in the interactive initialization of the program if automatic scaling is selected.

If the input data are in $f(X,Z)$ form, the cross-plot option prompt is displayed; otherwise the program proceeds to the "option menu." (See section below.) If the cross-plot option is selected, prompts for the Z-axis label, the Z-values, and the order of polynomial fit for the cross plot appear. After the user responds to these prompts, the program proceeds to the option menu.

Option Menu

The FIT option menu (fig. 2) is the central control point for the program from the user's perspective. A "HELP" response is available which explains each option in more detail. All program options are always displayed, although two, option <I> and option <Z>, are only available when the cross-plot option is activated.

Option <C> is selected to change the order of the polynomial fit, the attempted constraint setup, or the cross-plot option (if applicable). If the cross-plot option has been activated at any point during the terminal session, a prompt allows the user to switch the option on or off as desired. If

the cross-plot option has not yet been activated and the input data are in $X,f(X,Z)$ form, prompts requesting a Z-axis label, the Z-values, and a cross-plot order of the polynomial fit are issued.

Changing the axis scale setup, the plot options, the axis labels, or the plot title is accomplished by selecting option <S>. First, the current status of the program parameters which may be changed within this subroutine is displayed. Axis-scale-change options are then prompted. These options allow switching between manual scaling and automatic scaling or between different manual scaling schemes. A second X-axis may be added, or an existing one may be changed or removed. Data reference lines may be changed, added, or removed. Recall that the existence of a data reference line on an axis precludes the existence of a zero line or the plot grid. Zero lines may be added (if there is no data reference line) or removed. Recall also that the existence of a zero line precludes the existence of the plot grid. The plot grid may be added (if there is no reference line of any kind) or removed. The plot title may also be changed. Note that if this option is selected, the entire title must be entered. Then options to change the axis labels are prompted. As with the plot title, those labels selected for changes require that the entire label be entered.

Selecting option <E> to change the input data for the polynomial fit calls the "INPUT DATA EDITOR," which has its own submenu. All input data values may be changed, deleted, or added to. The INPUT DATA EDITOR contains extensive "escape" mechanisms to help the user avoid time-consuming mistakes. Entering a zero in response to any prompt requesting a number at a minimum escapes the user to the next prompt and no action is taken. For changes to input data, the user is always asked after completion of the changes if they are correct. If the changes are not correct, the program goes back to the beginning of the appropriate change section. When adding or deleting, the program prompts the user to provide the information to add or delete the appropriate points, but this information is not stored until approved by the user in response to a prompt.

It is sometimes useful to be able to compute and display the value of Y at a user-specified X-value (or X- and Z-values). Menu option <F> prompts the user to enter either an X-value or X- and Z-values (as appropriate). The Y-value is computed with the least-squares polynomial equation or equations, and then this value is printed on the screen. If the selected "point" is outside the range of the plotted polynomial fit, a warning message is printed to remind the user that although the polynomial fit may be accurate over the plotting range, the fit may diverge sharply outside of the range. This is especially true for the higher order polynomial fits.

With so many possible program parameters it is often desirable for the user to be able to review the current "position" of the various options. Option <D> prints every program parameter that the user can control and also gives the user the option to review the input data for the polynomial fit.

One graphic tool provided within FIT to help evaluate a cross-plot polynomial fit is called the "INTERMEDIATE Z-VALUES PLOT." This tool may be accessed by selecting option <I>. An "intermediate Z-value" is any Z-value within the range of all input Z-values. The primary usefulness of an intermediate Z-value plot is that it allows the user to check the desirability of a cross-plot polynomial fit by plotting Z-value contours between the contours of the polynomial fit through the data normally plotted. (See fig. 3.) Option <I> prompts the user to input the number of intermediate Z-values to be plotted (limit of 11, and selecting zero "escapes" the user back to the option menu) and then their values. After plotting the intermediate Z-values, a prompt asks if the user wishes to view more intermediate Z-value plots. If affirmative, the process is begun anew; otherwise, control is returned to the option menu.

The other graphic tool provided within FIT for use in analyzing the cross-plot polynomial fit is accessed by selecting option <Z>. Option <Z> is used to set up a plot of Y as a function of Z with X fixed. As shown in figure 4, the user gets a plot of Y in the plane of Z passing through a specified X-value. This is not quite analogous to the standard plot, which is a plot of Y in the plane of X passing through a set of specified Z-values. The usefulness of this plot of Z versus $f(X,Z)$ at a given X-value is that it provides an actual view of the surface fit in the Z-direction. When option <Z> is selected, the program computes the possible range of X-values and prints them before prompting the user for the X-value at which the plot is to be made. This response is checked to assure that an X-value has been selected which corresponds to the input X-value data. If no match occurs, the data X-values are printed and the prompt is repeated. Once a successful match occurs, Z versus $f(X,Z)$ is plotted. This X-value match is required so that the input data can be plotted as a reference for the polynomial fit. The user is then prompted for another plot or for return to the option menu.

Option <P> initiates the standard plotting sequence: Data which are in $X,f(X)$ form are plotted as such. If the data are in $X,f(X,Z)$ format, X versus $f(X,Z)$ is plotted, followed by the cross plot (if activated). (See figs. 5 and 6.) All the plotting options within FIT share a common plotting subroutine. The plotting subroutine makes extensive use of control flags to determine the logic path through it. The plot title is always written at the top of the plot. A subheader which gives the appropriate one of the three messages shown below is written under the title.

1. ORDER OF POLYNOMIAL FIT =
2. PLOT-DATA/NO-CURVE-FIT MODE
3. X( ) =

Axis scaling is then processed to define the appropriate data array locations and variables for use by the plotting routines. The axes are then plotted and labeled. If appropriate, data reference line (or lines), zero line (or lines), or the plot grid

are plotted. Then, based on the program settings, control is passed to one of four segments: the regular plot, the intermediate Z-value plot, the cross plot, or the $Z,f(X,Z)$ plot. When plotting is completed, the program waits, with no further action being taken until the user depresses the return key.

Output

FIT creates a variety of output files for potential use by the user. At the termination of the program, the program determines which files have been created and writes this information to the screen for the user. The two conditions checked by the program to determine what files have been created are whether a polynomial fit has been computed and whether the cross-plot option has been activated. The file DATIN is always created and is written in the format required by FIT to initialize the program. (See section entitled *Initialization of Program.*) If a polynomial fit has been computed, files DATOUT and DATCOEF are created. DATCOEF contains the "raw" coefficients of the polynomial fit in ascending order. DATOUT contains informative output in two sections for each polynomial fit computed during a terminal session. If the cross-plot option is activated, file DATCRSS is created and contains the raw coefficients from the cross-plot polynomial fit.

The informative output from the program is divided into two sections and is written on file DATOUT (as indicated above). Each time a polynomial fit is computed, output is written to the DATOUT file. The first section provides the following information:

1. Scale factors applied to the input data (see *Input Data Processing* section)
2. A complete list of the data points being used for the polynomial fit
3. Data points that are weighted (attempted constraint) for the polynomial fit, denoted by an asterisk
4. At each data point, the computed value of Y
5. At each data point, the error between the input data Y-value and the computed value of Y
6. At each data point, the appropriate first derivative or derivatives
7. A summary listing of average error, average percent error, root-mean-square (rms) error, and the sum of the residuals from the polynomial fit

The second section of informative output is the polynomial fit coefficients written in "equation format." The equation format shows how the polynomial equations should be written to duplicate the plotted polynomial fit. This feature is especially important and useful when using the cross-plot option, since application of the output coefficients in the proper equations is critical.

Error Trapping and Help

Throughout the interactive subroutines within FIT, user responses are checked whenever improper inputs could potentially cause a program error. The program uses 19 informative messages which are selected as appropriate to inform the user of the "error" in the input response. Some messages are warnings; most flag the error made and ask the user to try again. Other messages inform the user of problems and describe action that the program has taken to resolve them. The program contains additional information concerning the cross plot and the attempted constraint of the polynomial fit at certain data points. The program also contains general information about the allowable range in the input data values and the selection of the order of the polynomial fit. The appropriate information is displayed whenever "HELP" is selected in response to a yes/no prompt.

Examples

Sample input for the program is shown in appendixes A, C, and D. Appendix A is the DATIN file created during a program run. If DATIN exists in the local file environment when FIT is executed, the program reads the file as described previously. Appendix D shows the screen prompts and user responses from a terminal session. Figure 5 shows the plot generated from these data. Appendix E is the DATOUT file of informative output created during this program run. Appendix C is similar to appendix D except the input data for the polynomial fit are read from a file, the data are in $X,f(X,Z)$ format, and the cross plot is activated. Figure 6 shows the plots generated from these data. Figures 3 and 4 show the prompts, user responses, and plots generated when options <I> and <Z>, respectively, are selected and the input from appendix C is used. The DATOUT file from an appendix C program run is shown in appendix F.

Concluding Remarks

FIT is a useful interactive computer program for developing polynomial fits of data that describe a curve or a surface. The program includes a unique methodology for computing the equation of a polynomial fit through a surface. Computer graphics are provided with the program to allow the engineer to visualize the computed fit and to better interact with the program capabilities. These capabilities allow the engineer to determine polynomial equations fitted through data for most realistic applications.

NASA Langley Research Center
Hampton, VA 23665
May 21, 1985

References

1. Lawson, Charles L.; and Hanson, Richard J.: *Solving Least Squares Problems*. Prentice-Hall, Inc., 1974.

2. Businger, Peter; and Golub, Gene H.: Linear Least Squares Solutions by Householder Transformations. *Numer. Math.*, Bd. 7, Heft 3, 1965, pp. 269–276. Reprinted in *Handbook for Automatic Computation, Volume II—Linear Algebra*, J. H. Wilkinson and C. Reinsch, eds., Springer-Verlag, 1971, pp. 111–118.

Appendix A

Sample DATIN File for Input From Appendix D

```
1
                                    APPENDIX A/D
                                      X
                                F(X)
6 1 3 1. 1.
0. 1.
1. .8629504
2. .5360412
3. .5761378
4. .4456731
5. .4142661
0
1 0 0 0 0 0
.2 0. 1.2 2.
0
0
```

Appendix B

Input Data Format for FIT

The following data are read by rows.

X	$f(X,Z_1)$	$f(X,Z_2)$	$f(X,Z_3)$	$f(X,Z_4)$	$f(X,Z_5)$
0.4	0.270	0.165	0.060	−0.040	−0.135
0.6	0.145	0.070	0.010	−0.065	−0.140
0.8	0.085	0.035	−0.020	−0.075	−0.125
1.0	0.060	0.020	−0.025	−0.070	−0.115
1.2	0.040	0.010	−0.030	−0.065	−0.100
1.4	0.030	0	−0.030	−0.060	−0.090
1.6	0.025	−0.005	−0.030	−0.055	−0.080
1.8	0.020	−0.010	−0.030	−0.050	−0.075
2.0	0.017	−0.015	−0.030	−0.045	−0.072

Appendix C

Typical Screen Prompts and User Responses During Surface Fit Terminal Session

```
                    **** WELCOME TO FIT ****
      ** VERSION 5.6  --  LAST UPDATED OCTOBER 01, 1984 **


    TYPE "HELP" FOR AN EXPLANATION OF FIT, HIT CR TO CONTINUE ?
    OPTIONS FOR INPUTTING THE X AND CORRESPONDING F(X,Z) VALUES:
        0= READ X-F(X,Z) DATA FROM INPUT KEYED IN AT TERMINAL
        1= READ X-F(X,Z) DATA FROM ONE OF YOUR LOCAL FILES
        2= EXIT FIT TO MAKE PERMANENT FILE(S) LOCAL

    SELECT OPTION --> ? 1
    ENTER YOUR LOCAL FILE NAME TO BE READ ? FITDAT
    ENTER NUMBER OF INDEPENDENT DATA POINTS (X-VALUES ; LIMIT 100)  ? 9
    ENTER NUMBER OF DEPENDENT DATA POINTS (F(X,Z) VALUES) PER X-VALUE
    (LIMIT 11)  (FOR EXPLANATION OF THIS PROMPT TYPE "HELP")  ? 5
    ENTER PLOT TITLE (60 SPACES)
    1        10        20        30        40        50        60
    *        *         *         *         *         *         *
  ?                          APPENDIX C

    ENTER X-LABEL (50 SPACES)
                              X
    ENTER F(X,Z) LABEL (40 SPACES)
                     F(X,Z)
    ENTER ORDER OF LEAST SQUARES FIT  (LIMIT 7) ? 5
    DO YOU WANT TO ATTEMPT TO CONSTRAIN THE CURVE FIT(S) TO GO THROUGH
    CERTAIN POINTS (Y/N/HELP/RET=HELP) ? N
    IS AUTOMATIC DATA SCALING DESIRED ON THE X-AXIS (Y/N/RET=Y) ? Y
    IS AUTOMATIC DATA SCALING DESIRED ON F(X,Z) AXIS (Y/N/RET=Y)  ? Y
    PLOT DATA REFERENCE LINES (Y/N/RET=N) ? N
    PLOT HORIZONTAL "ZERO LINE" (Y/N/RET=N) ? N
    DRAW PLOT GRID (Y/N/RET=N) ? N
    IS A CROSS PLOT OF THE COEFFICIENTS DESIRED (Y/N/HELP/RET=HELP) ? Y
    ENTER Z-AXIS LABEL (HORZ. AXIS ON CROSS PLOT ; 50 SPACES)
    1        10        20        30        40        50
    *        *         *         *         *         *
                              Z
    ENTER 5 Z-VALUES (VALUES OF CONSTANT Z FROM FIRST PLOT TO BE USED
    IN CROSS PLOT)
  ? 0 5 10 15 20
    ENTER ORDER OF CROSS PLOT FIT (LIMIT 4) ? 4

    IF COPY OF INPUT IS DESIRED, MAKE IT NOW.

    HIT RETURN TO CONTINUE
      (SCREEN WILL CLEAR) ?
```

Appendix D

Typical Screen Prompts and User Responses During Curve Fit Terminal Session

```
                    **** WELCOME TO FIT ****
          ** VERSION 5.6  --  LAST UPDATED OCTOBER 01, 1984 **


  TYPE "HELP" FOR AN EXPLANATION OF FIT, HIT RETURN TO CONTINUE ?
  OPTIONS FOR INPUTTING THE X AND CORRESPONDING F(X,Z) VALUES:
      0= READ X-F(X,Z) DATA FROM INPUT KEYED IN AT TERMINAL
      1= READ X-F(X,Z) DATA FROM ONE OF YOUR LOCAL FILES
      2= EXIT FIT TO MAKE PERMANENT FILE(S) LOCAL

  SELECT OPTION --> ? 0
  ENTER NUMBER OF INDEPENDENT DATA POINTS (X-VALUES ; LIMIT 100)  ? 6
  ENTER NUMBER OF DEPENDENT DATA POINTS (F(X,Z) VALUES) PER X-VALUE
  (LIMIT 11)  (FOR EXPLANATION OF THIS PROMPT TYPE "HELP")  ? 1
  ENTER NUMBER 1 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 0 1
  ENTER NUMBER 2 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 1 .8629504
  ENTER NUMBER 3 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 2 .5360412
  ENTER NUMBER 4 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 3 .5761378
  ENTER NUMBER 5 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 4 .4456731
  ENTER NUMBER 6 VALUE OF X AND 1 CORRESPONDING  F(X) -VALUE(S)
? 5 .4142661
  ENTER PLOT TITLE (60 SPACES)
  1       10        20        30        40        50        60
  *        *         *         *         *         *         *
?                          APPENDIX A/D

  ENTER X-LABEL (50 SPACES)
?                          X
  ENTER  F(X)  LABEL (40 SPACES)
?                 F(X)
  ENTER ORDER OF POLYNOMIAL FIT  (LIMIT 5)  ? 3
  DO YOU WANT TO ATTEMPT TO CONSTRAIN THE CURVE FIT(S) TO GO THROUGH
  CERTAIN POINTS (Y/N/HELP/RET=HELP)? N
  IS AUTOMATIC DATA SCALING DESIRED ON THE X-AXIS (Y/N/RET=Y) ?
  IS AUTOMATIC DATA SCALING DESIRED ON  F(X)  AXIS (Y/N/RET=Y)  ? N

   F(X)  VALUES RANGE FROM .4142661 TO 1.

   ** FOLLOWING 3 PROMPTS PERTAIN TO  F(X)  AXIS **
    ENTER  F(X)  ORIGIN VALUE ? 0
    ENTER MAXIMUM  F(X)  VALUE ON AXIS  ? 1.2
  ENTER DESIRED NUMBER OF MAJOR TIC MARKS (RET=8) ? 6
```

Appendix D

Typical Screen Prompts and User Responses During Curve Fit Terminal Session (concluded)

```
PLOT DATA REFERENCE LINES (Y/N/RET=N) ?
DRAW PLOT GRID (Y/N/RET=N)?
IS A SECOND X-AXIS DESIRED (Y/N/RET=N)  ?

IF COPY OF INPUT IS DESIRED, MAKE IT NOW.

HIT RETURN TO CONTINUE
  (SCREEN WILL CLEAR) ?
```

Appendix E

DATOUT File of Informative Output Corresponding to the Input of Appendixes A and D

```
***   PROGRAM FIT VERSION 5.6 INFORMATIVE OUTPUT   ***

                APPENDIX A/D

            ORDER OF POLYNOMIAL FIT=  3

POINT X-VALUE CURVE    X-       F(X)-     F(X)-      ERROR        DF/DX
 NO.    NO.    NO.    DATA      DATA      CALC.
-----  -------  -----  ----      ----      -----      -----        -----

  1       1       1    .0000   1.0000   1.0178  .1780E-01 -.2569
  2       2       1   1.0000    .8630    .7944 -.6851E-01 -.1913
  3       3       1   2.0000    .5360    .6320  .9598E-01 -.1351
  4       4       1   3.0000    .5761    .5212 -.5494E-01 -.8814E-01
  5       5       1   4.0000    .4457    .4526  .6954E-02 -.5055E-01
  6       6       1   5.0000    .4143    .4170  .2713E-02 -.2230E-01

          AVERAGE ERROR=          .411501E-01
          AVE. PERCENT ERROR=     6.56266
          RMS ERROR=              .344882E-01

  SUM OF RESIDUALS (FROM CURVE FIT ROUTINE) =     .1730E-01


POLYNOMIAL EQUATION FORM WITH COEFFICIENTS FROM CURVE FITTING ROUTINE ARE WRITTEN BELOW

                ORDER OF POLYNOMIAL =  3

X DENOTES THE INDEPENDENT VARIABLE, F(X) DENOTES THE DEPENDENT VARIABLE

F(X)= ( 1.01780    )X**0  + (-.256944    )X**1  + ( .351388E-01)X**2  + (-.155662E-02)X**3
```

Appendix F

DATOUT File of Informative Output Corresponding to the Input of Appendix C

```
***   PROGRAM FIT VERSION 5.6 INFORMATIVE OUTPUT   ***

                APPENDIX C

          ORDER OF POLYNOMIAL FITS:
             FIRST PLOT =  5
             CROSS PLOT =  4

        X-DATA SCALE FACTOR =  1.00000
        F(X,Z)-DATA SCALE FACTOR = .100000
```

PT NO.	X-VAL NO.	Z-VAL NO.	SCALED X- DATA	UNSCALED Z- DATA	SCALED F(X,Z)- DATA	SCALED F(X,Z)- CALC.	UNSCALED ERROR	UNSCALED DF/DX	UNSCALED DF/DZ
1	1	1	.4000	.0000	2.7000	2.7000	-.2331E-05	-.8796	-.2061E-01
2	1	2	.4000	5.0000	1.6500	1.6490	-.1049E-03	-.7120	-.2117E-01
3	1	3	.4000	10.0000	.6000	.6030	.3030E-03	-.3462	-.2054E-01
4	1	4	.4000	15.0000	-.4000	-.3979	.2098E-03	-.1990	-.1950E-01
5	1	5	.4000	20.0000	-1.3500	-1.3530	-.3007E-03	-.7976E-01	-.1883E-01
6	2	1	.6000	.0000	1.4500	1.4480	-.2034E-03	-.4211	-.1975E-01
7	2	2	.6000	5.0000	.7000	.7043	.4312E-03	-.2852	-.1214E-01
8	2	3	.6000	10.0000	.1000	.0848	-.1518E-02	-.1834	-.1347E-01
9	2	4	.6000	15.0000	-.6500	-.6601	-.1005E-02	-.7429E-01	-.1584E-01
10	2	5	.6000	20.0000	-1.4000	-1.3855	.1448E-02	.3148E-01	-.1138E-01
11	3	1	.8000	.0000	.8500	.8609	.1087E-02	-.1955	-.1118E-01
12	3	2	.8000	5.0000	.3500	.3446	-.5361E-03	-.1047	-.9989E-02
13	3	3	.8000	10.0000	-.2000	-.1714	.2861E-02	-.8166E-01	-.1080E-01
14	3	4	.8000	15.0000	-.7500	-.7327	.1731E-02	-.6171E-02	-.1144E-01
15	3	5	.8000	20.0000	-1.2500	-1.2756	-.2562E-02	.6957E-01	-.9705E-02
16	4	1	1.0000	.0000	.6000	.5788	-.2123E-02	-.1016	-.5013E-02
17	4	2	1.0000	5.0000	.2000	.2003	.3497E-04	-.5325E-01	-.9195E-02
18	4	3	1.0000	10.0000	-.2500	-.2726	-.2258E-02	-.2604E-01	-.9290E-02
19	4	4	1.0000	15.0000	-.7000	-.7102	-.1023E-02	.2361E-01	-.8280E-02
20	4	5	1.0000	20.0000	-1.1500	-1.1319	.1811E-02	.7034E-01	-.9148E-02
21	5	1	1.2000	.0000	.4000	.4159	.1585E-02	-.6652E-01	-.3823E-02
22	5	2	1.2000	5.0000	.1000	.1012	.1166E-03	-.4871E-01	-.7814E-02
23	5	3	1.2000	10.0000	-.3000	-.2971	.2914E-03	-.2742E-02	-.7698E-02
24	5	4	1.2000	15.0000	-.6500	-.6535	-.3497E-03	.3041E-01	-.6673E-02
25	5	5	1.2000	20.0000	-1.0000	-1.0024	-.2448E-03	.5846E-01	-.7933E-02
26	6	1	1.4000	.0000	.3000	.3039	.3899E-03	-.4556E-01	-.5719E-02
27	6	2	1.4000	5.0000	.0000	.0069	.6876E-03	-.4329E-01	-.6068E-02
28	6	3	1.4000	10.0000	-.3000	-.2965	.3526E-03	.1141E-02	-.6030E-02
29	6	4	1.4000	15.0000	-.6000	-.5955	.4458E-03	.2667E-01	-.5955E-02
30	6	5	1.4000	20.0000	-.9000	-.8972	.2774E-03	.4744E-01	-.6193E-02
31	7	1	1.6000	.0000	.2500	.2357	-.1425E-02	-.2234E-01	-.7469E-02
32	7	2	1.6000	5.0000	-.0500	-.0619	-.1189E-02	-.2388E-01	-.4949E-02
33	7	3	1.6000	10.0000	-.3000	-.2975	.2506E-03	-.2465E-02	-.4730E-02
34	7	4	1.6000	15.0000	-.5500	-.5474	.2622E-03	.2198E-01	-.5261E-02
35	7	5	1.6000	20.0000	-.8000	-.8103	-.1028E-02	.3965E-01	-.4992E-02
36	8	1	1.8000	.0000	.2000	.2087	.8735E-03	-.8753E-02	-.7612E-02
37	8	2	1.8000	5.0000	-.1000	-.0929	.7110E-03	-.1199E-01	-.4815E-02
38	8	3	1.8000	10.0000	-.3000	-.3040	-.3992E-03	-.2589E-02	-.3894E-02
39	8	4	1.8000	15.0000	-.5000	-.5038	-.3759E-03	.2300E-01	-.4264E-02
40	8	5	1.8000	20.0000	-.7500	-.7421	.7902E-03	.2628E-01	-.5339E-02
41	9	1	2.0000	.0000	.1700	.1682	-.1818E-03	-.4493E-01	-.9561E-02
42	9	2	2.0000	5.0000	-.1500	-.1515	-.1515E-03	-.6374E-01	-.3994E-02
43	9	3	2.0000	10.0000	-.3000	-.2988	.1166E-03	.1078E-01	-.2482E-02
44	9	4	2.0000	15.0000	-.4500	-.4490	.1049E-03	.3356E-01	-.3922E-02
45	9	5	2.0000	20.0000	-.7200	-.7219	-.1911E-03	-.1260E-01	-.7212E-02

```
        AVERAGE ERROR=          .763222E-03
        AVE. PERCENT ERROR=     2.66258
        RMS ERROR=              .723775E-03

   SUM OF RESIDUALS (FROM FIRST PLOT)   .1122E-02   .2914E-03   .1613E-02   .5638E-03   .1388E-02

   SUM OF RESIDUALS (FROM CROSS PLOT)   0.          0.          0.          0.          0.

   SUM OF RESIDUALS (FROM CROSS PLOT)   0.

COFFFICIENT NORMALIZING FACTORS=
        .4465       1.6662       2.3509       1.6597       .5829

        .0811
```

Appendix F

DATOUT File of Informative Output Corresponding to the Input of Appendix C (concluded)

```
     POLYNOMIAL EQUATION FORM WITH COEFFICIENTS FROM CURVE FITTING ROUTINE ARE WRITTEN BELOW

                                   ORDER OF POLYNOMIAL =  5

 X DENOTES THE INDEPENDENT VARIABLE, F(X,Z) DENOTES THE DEPENDENT VARIABLE OF THE FIRST PLOT, AND Z DENOTES
 THE INDEPENDENT VARIABLE OF THE CROSS PLOT


  C( 1)= ( .984833     )Z**0  + ( .977694E-01)Z**1  + (-.391192E-01)Z**2  + ( .291189E-02)Z**3  + (-.661667E-04)Z**4
  C( 2)= (-3.04160     )Z**0  + (-.689331    )Z**1  + ( .211062    )Z**2  + (-.156405E-01)Z**3  + ( .355043E-03)Z**4
  C( 3)= ( 4.15186     )Z**0  + ( 1.41104    )Z**1  + (-.403061    )Z**2  + ( .297466E-01)Z**3  + (-.673866E-03)Z**4
  C( 4)= (-2.93238     )Z**0  + (-1.26433    )Z**1  + ( .351132    )Z**2  + (-.258434E-01)Z**3  + ( .584286E-03)Z**4
  C( 5)= ( 1.04240     )Z**0  + ( .520824    )Z**1  + (-.142783    )Z**2  + ( .104891E-01)Z**3  + (-.236682E-03)Z**4
  C( 6)= (-.147236     )Z**0  + (-.809796E-01)Z**1  + ( .220469E-01)Z**2  + (-.161859E-02)Z**3  + ( .363916E-04)Z**4

F(X,Z)= C(1)X**0  +C(2)X**1  +C(3)X**2  +C(4)X**3  +C(5)X**4  +C(6)X**5
```

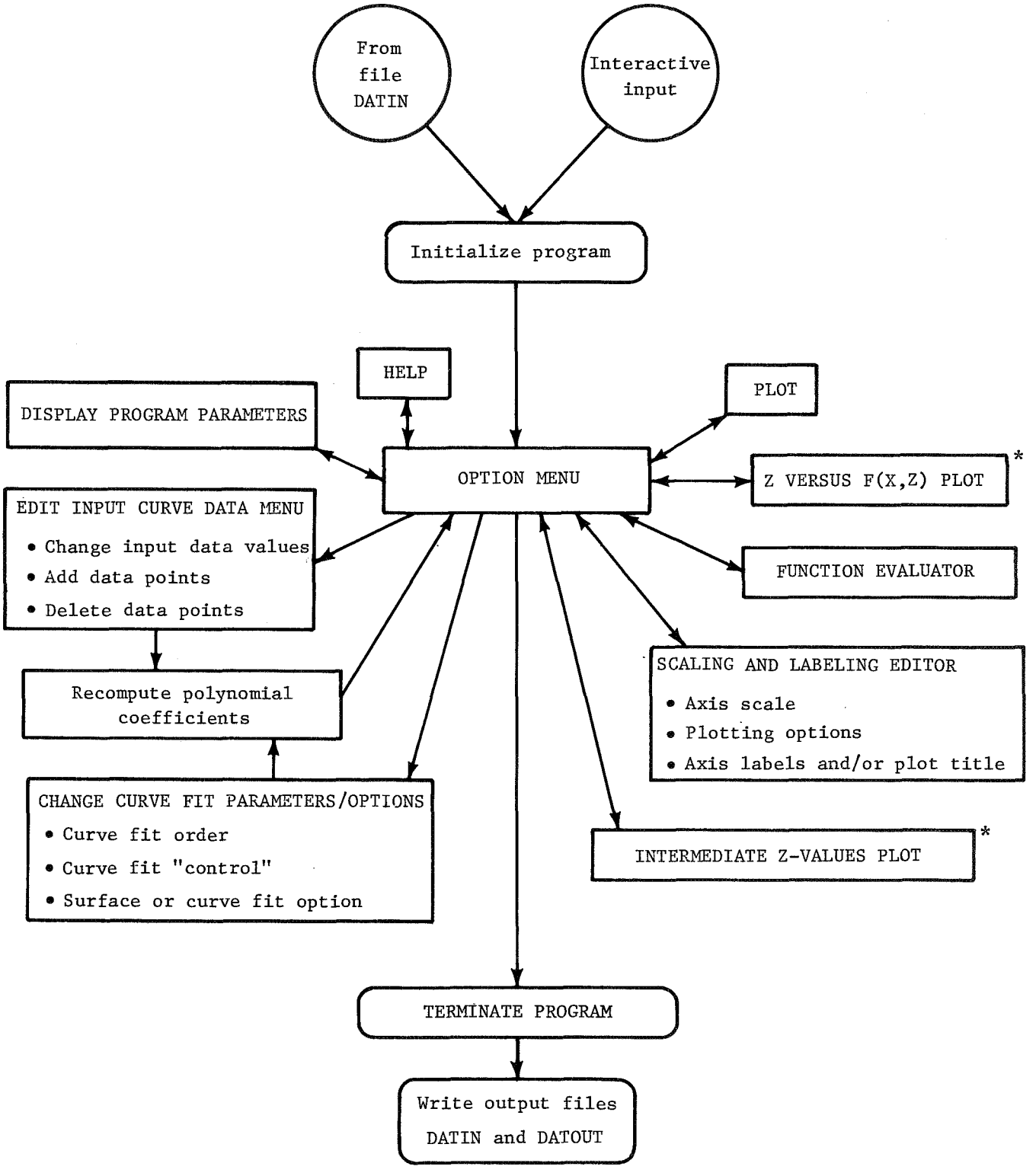
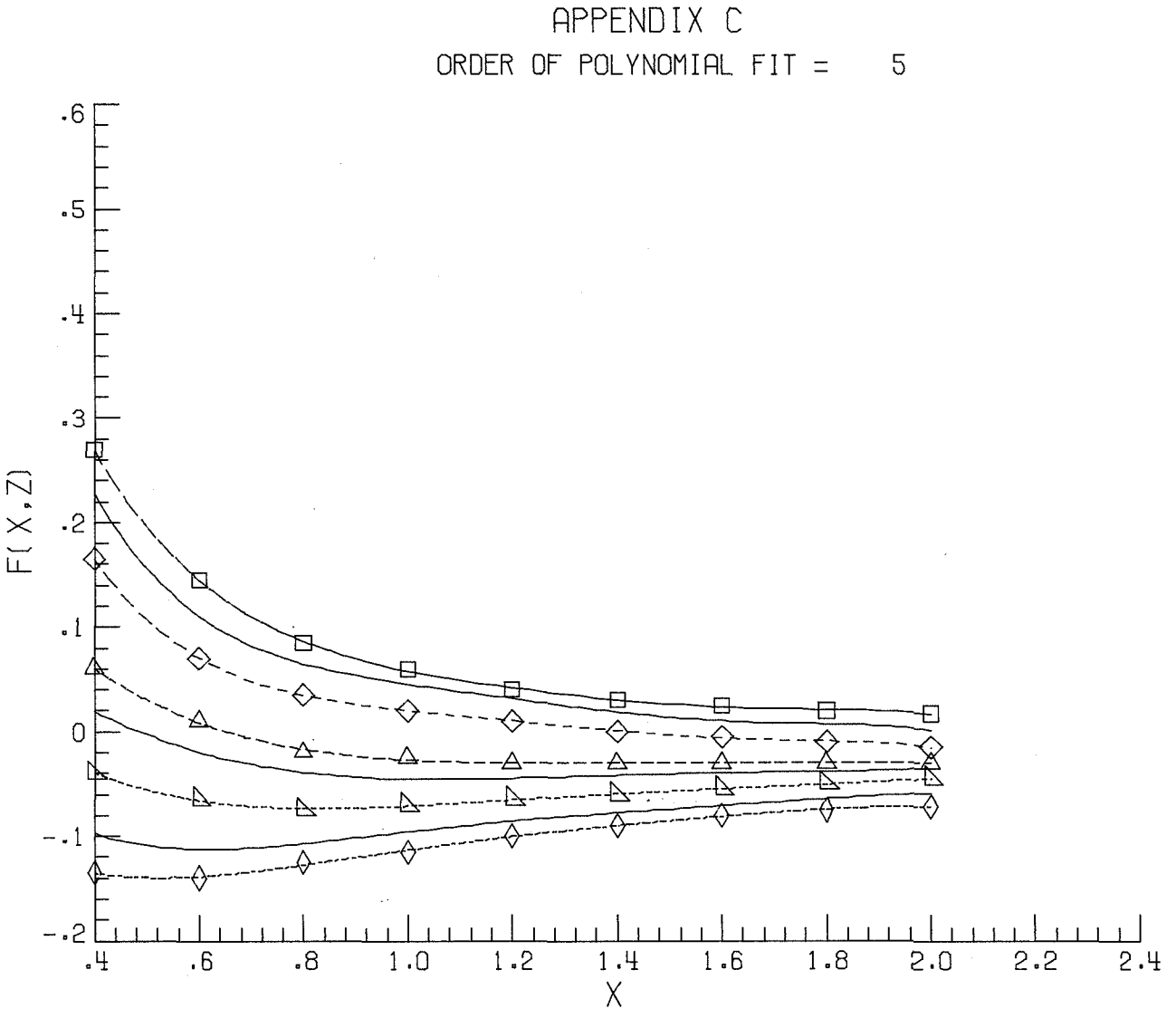


Figure 1. - FIT program hierarchy. Asterisk denotes options that are only available when surface fitting.

```
   ***    PROGRAM FIT VERSION 5.6    ***

        >>>       OPTION MENU       <<<


      <C>HANGE CURVE FIT PARAMETERS/OPTIONS
      <D>ISPLAY PROGRAM PARAMETERS
      <E>DIT INPUT CURVE DATA
      <F>UNCTION EVALUATOR
      <I>NTERMEDIATE Z-VALUES PLOT
      <P>LOT
      <S>CALING AND LABELING EDITOR
      <T>ERMINATE PROGRAM
      <Z> VERSUS F(X,Z) PLOT

***   <H>ELP


 SELECT OPTION --> ?
```

Figure 2. - FIT option menu.

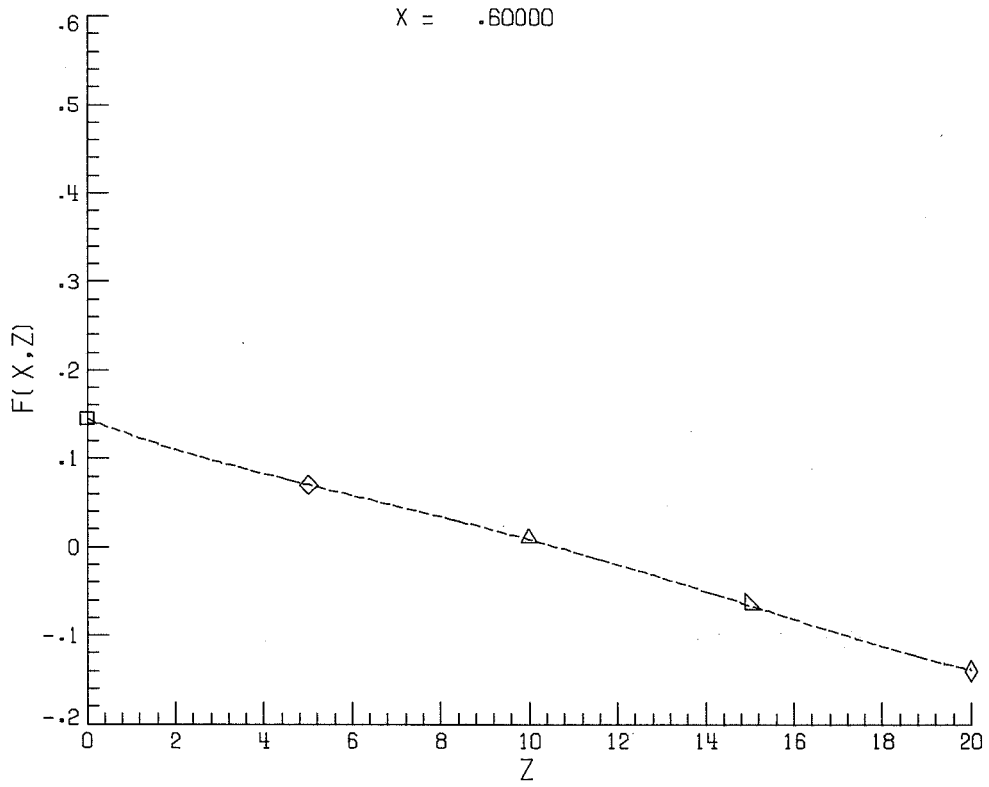
```
HOW MANY INTERMEDIATE VALUES DO YOU WANT TO PLOT (LIMIT 11) ? 3
ENTERED VALUES MAY RANGE FROM 0. TO 20.
ENTER 3 VALUES TO BE PLOTTED
? 2 12 18
THE CURVE FIT WILL BE PLOTTED FOR THE FOLLOWING Z-VALUES:
  Z = 2.000
  Z = 12.00
  Z = 18.00
CORRECT SET UP (Y/N/RET=Y)  ?
```




```
DO YOU WANT TO PLOT MORE INTERMEDIATE VALUES OF
                              Z                          (Y/N/RET=N)  ?
```

Figure 3. - Screen prompts, user responses, and plot when option <I> is selected.

```
 ENTERED VALUE MAY RANGE FROM .4 TO 2.
 INPUT X-VALUE AT WHICH  Z,F(X,Z)  PLOT IS TO BE MADE
(RET=ESCAPE TO MAIN MENU)  ? .55

 NO X-VALUE MATCH -- X-VALUES LISTED BELOW

 X(  1)=  .4000            X(  2)=  .6000
 X(  3)=  .8000            X(  4)=  1.000
 X(  5)=  1.200            X(  6)=  1.400
 X(  7)=  1.600            X(  8)=  1.800
 X(  9)=  2.000

HIT RETURN TO CONTINUE
 (SCREEN WILL CLEAR)
?


 ENTERED VALUE MAY RANGE FROM .4 TO 2.
 INPUT X-VALUE AT WHICH  Z,F(X,Z)  PLOT IS TO BE MADE
(RET=ESCAPE TO MAIN MENU)  ? .6
   PLOT WILL BE DRAWN AT X =  .6000
 CORRECT SET UP (Y/N/RET=Y)  ?
```

FIT 5.6

```
DO YOU WANT TO PLOT ANOTHER  Z,F(X,Z)  PLOT (Y/N/RET=N) ?
```

Figure 4. - Screen prompts, user responses, and plot when option <Z> is selected.

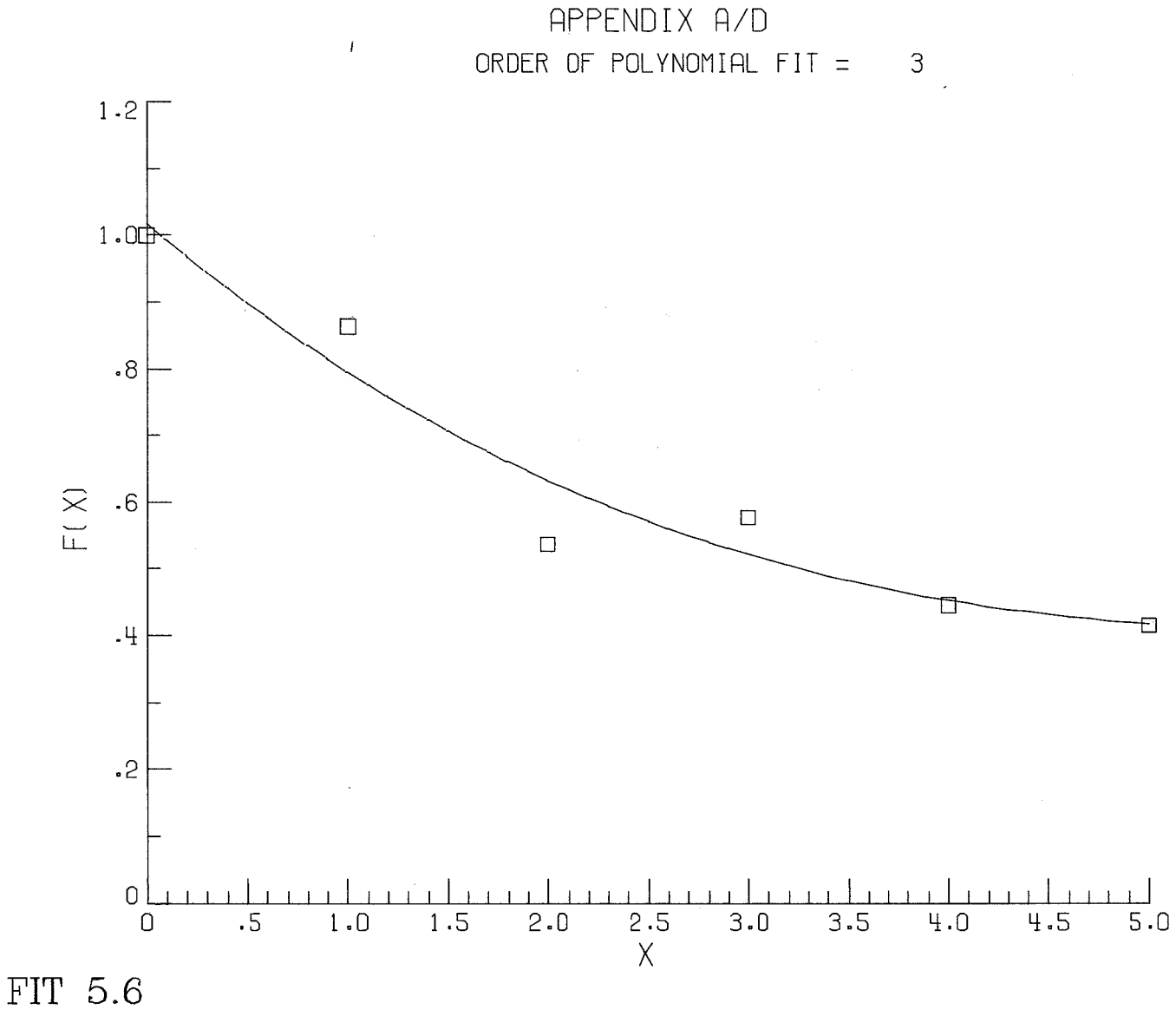
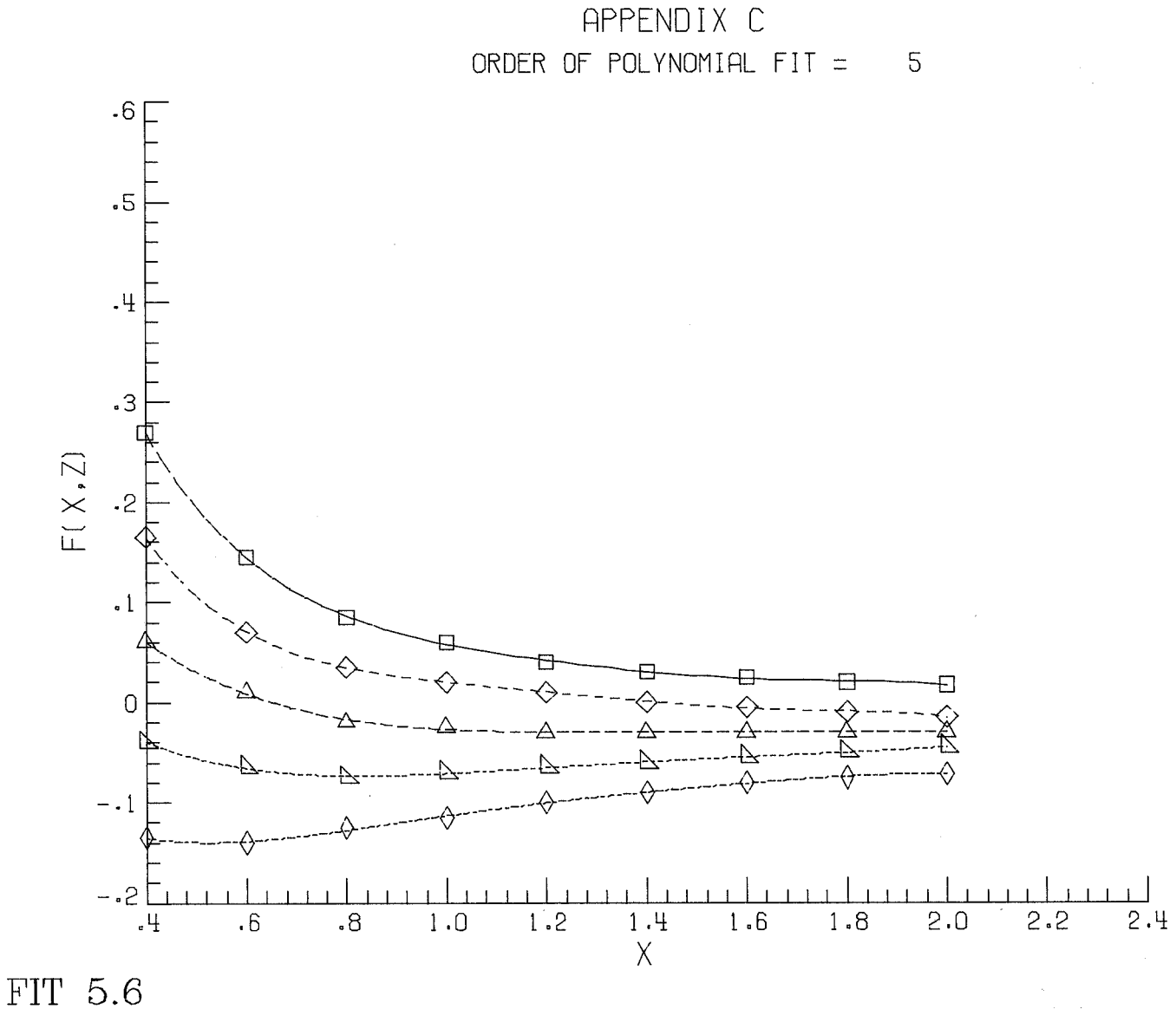
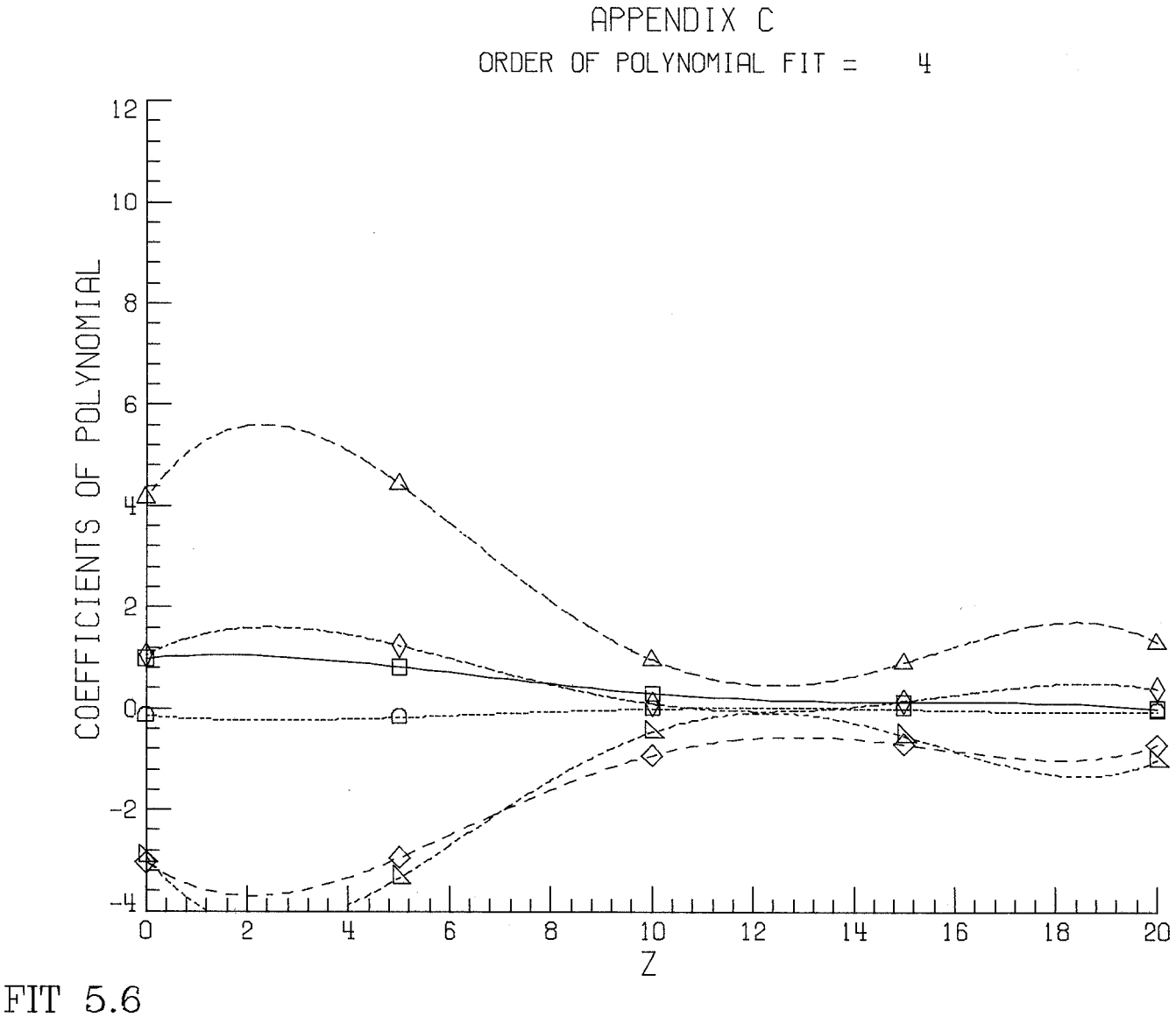


Figure 5. - Plot generated from input shown in appendixes A and D.



(a) Regular plot.

Figure 6. - Plot generated from input shown in appendix C.



(b) Cross plot.

Figure 6. - Concluded.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM-86413		

4. Title and Subtitle	5. Report Date
FIT: Computer Program That Interactively Determines Polynomial Equations for Data Which Are a Function of Two Independent Variables	October 1985
	6. Performing Organization Code 505-34-03-05
7. Author(s) P. Douglas Arbuckle, Steven M. Sliwa, Marie-Louise Roy, and Sherwood H. Tiffany	8. Performing Organization Report No. L-15983
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665	10. Work Unit No.
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546	13. Type of Report and Period Covered Technical Memorandum
	14. Sponsoring Agency Code

15. Supplementary Notes

P. Douglas Arbuckle, Steven M. Sliwa, and Sherwood H. Tiffany: Langley Research Center, Hampton, Virginia.
Marie-Louise Roy: PRC Kentron, Inc., Hampton, Virginia.

16. Abstract

A computer program for interactively developing least-squares polynomial equations to fit user-supplied data is described. The program is characterized by the ability to compute the polynomial equations of a "surface" fit through data that are a function of two independent variables. The program utilizes the Langley Research Center graphics packages to display polynomial equation curves and data points, facilitating a qualitative evaluation of the effectiveness of the fit. An explanation of the fundamental principles and features of the program, as well as sample input and corresponding output, are included.

17. Key Words (Suggested by Authors(s))	18. Distribution Statement
Interactive computer program Polynomial fit Independent variables Polynomial equation coefficients Curve fitting Surface fitting	Unclassified—Unlimited Subject Category 08

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	19	A02

For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA-Langley, 1985